Quality of Experience Assessment of Video in Real-life Environments

Nicolas Staelens

Supervisor(s): Brecht Vermeulen, Piet Demeester

I. Introduction

According to ITU-T Rec. P.10/G.100 Amd. 2, Quality of Experience (QoE) includes the complete end-to-end system effects and can be influenced by user expectations and context. Therefore, the QoE of IPTV and Video on Demand (VoD) services should ideally be measured in the home environment. Unfortunately, current existing subjective video quality assessment methodologies require a controlled environment where human observers rate the visual quality of a series of short video clips (10 - 15 sec). In this article, we compare the results of a subjective test conducted in real-life environments with the results obtained using a standardized methodology.

II. Real-life Quality of Experience Assessment

Existing subjective quality assessment methodologies, specified in ITU-T Rec. P.910 and ITU-R Rec. BT.500-11, pose stringent demands on the way subjective experiments are set up and conducted. As a consequence, they cannot be used for the quality assessment of long video sequences in realistic environments.

In order to enable real-life QoE assessment of IPTV and VoD services, we used a novel methodology based on full length DVD movies. This new methodology tries to mimic the typical lean-backward TV-viewing experience by letting the subjects take the movie home to watch it in the same environment they usually watch television. Using the toolchain described in [1], controlled visual degradations can be inserted in the movie prior to writing it on a compliant DVD disc.

III. Subjective Test

Two subjective tests were conducted using our new methodology described in the previous section. In [1], we used movies which contained up to three visual impairments to study the tolerance towards random blockiness and frame freezes. A second subjective test was conducted in [2] in order to assess the influence of H.264 SVC scalability on end-users' QoE. During this test, a movie was used in which six fragments were temporally and quality (SNR) scaled down. In both tests, subjects were not informed about any possible degradations.Feedback was collected by means of a questionnaire which was provided to the subjects in a sealed envelope.

In order to highlight the importance of measuring QoE in real-life environments, we compared the results from [1] and [2] with the results obtained using a standardized assessment methodology[1]. For this standard test, we extracted the fragments containing a visual impairment from the DVDs in order to create short video sequences of 15 seconds long.

N. Staelens is with the Department of Information Technology, Ghent University - IBBT, Belgium. E-mail: nicolas.staelens@intec.ugent.be .

[1] based on the Single Stimulus Absolute Category Rating (SS ACR) method from ITU-R Rec. BT.500-11.

IV. Results

As shown in Figure 1, visual impairments are much more detected during a standard subjective test. Before such tests, subjects receive specific instructions on how to evaluate the video sequences. As a consequence, subjects' primary focus is on visual quality evaluation. While watching a movie, subjects are concentrated on content.

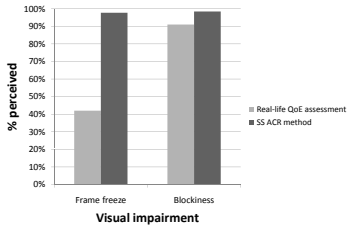


Figure 1. Impairment detection during real-life QoE assessment and a standard subjective test.

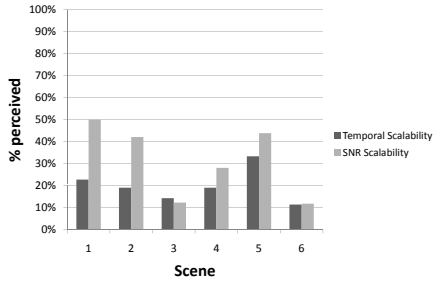
Despite the fact that frame freezes are less detected, they were rated more annoying during movies as they break the natural flow and the fluidity of the playback. However, during the standard test, quality scores for frame freezes were significantly higher compared to the scores for blockiness.

A similar result is obtained regarding the visibility of temporal and quality scalability, depicted in Figure 2. The graph indicates that impairment visibility is also content dependent.

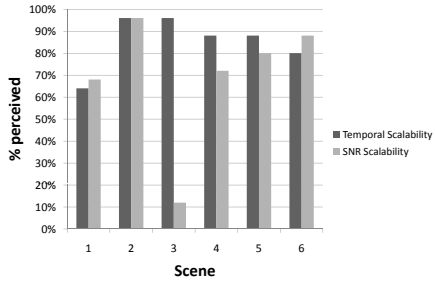
Overall quality ratings in [1] showed that subjects tolerate up to two visual impairments during playback. It is important to mention that this quality rating is higher compared to the ratings subjects gave to the short clips, which only contained one visual degradation.

V. Conclusions

Our results show significant differences concerning impairment visibility and tolerance when conducting subjective video quality tests in real-life environments. Furthermore, our research indicates that results obtained using a standard video quality assessment methodology do not always hold in the case of real-life QoE assessment.



(a)Real-life QoE assessment



(b)SS ACR method

Figure 2. Downscaling visibility during real-life assessment (a) and during a standard test (b).

Acknowledgments

This paper is the result of research carried out as part of the Video Q-SAC project, partly funded by the IBBT.

References

[1] N. Staelens, B. Vermeulen, S. Moens, J.-F. Macq, P. Lambert, R. Van de Walle, and P. Demeester, "Assessing the Influence of Packet Loss and Frame Freezes on the Perceptual Quality of Full Length Movies," *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, January 2009.

[2] N. Staelens, S. Moens, W. Van den Broeck, I. Mariën, B. Vermeulen, P. Lambert, R. Van de Walle, and P. Demeester, "Assessing the Perceptual Influence of H.264/SVC Signal-to-Noise Ratio and Temporal Scalability on Full Length Movies," *First International Workshop on Quality of Multimedia Experience*, July 2009.